



## Análisis de Datos



Jesús García Herrero

### **Práctica 2 WEKA. Utilización de filtros con WEKA.**

En este ejercicio se utilizarán facilidades de filtrado de atributos con WEKA. Cargar el fichero de datos “mushroom.arff”. Consiste en 8124 instancias de tipos de setas, con 22 atributos nominales y la clase: ‘e’, comestible (eatable), ‘p’, venenosa (poison).

1. Comparar el error de clasificación con validación cruzada utilizando los clasificadores “naive bayes” y “C4.5”.
2. Con los atributos seleccionados por el clasificador C4.5 (parámetros por defecto), incluir un filtro de atributos en el cuadro “Preprocess”. Es un filtro de tipo “no supervisado” de selección de índices (“Remove”), donde se indicarán los índices de los atributos seleccionados. Comparar el error alcanzado por el clasificador “naive bayes” en este caso.
3. Utilizar ahora un filtro de tipo “supervisado”. Para ello, desactivando cualquier filtro de “Preprocess”, utilizar un clasificador con filtro “FilteredClassifier”, donde el clasificador será el “naive bayes”, y el filtro de tipo supervisado con selección de atributos.